

Analysis-by-Synthesis Quantization for Compressed Sensing Measurements

Amirpasha Shirazinia, *Student Member, IEEE*, Saikat Chatterjee, *Member, IEEE*, Mikael Skoglund, *Senior Member, IEEE*

Abstract—We consider a resource-limited scenario where a sensor that uses compressed sensing (CS) collects a low number of measurements in order to observe a sparse signal, and the measurements are subsequently quantized at a low bit-rate followed by transmission or storage. For such a scenario, we design new algorithms for source coding with the objective of achieving good reconstruction performance of the sparse signal. Our approach is based on an analysis-by-synthesis principle at the encoder, consisting of two main steps: (1) the synthesis step uses a sparse signal reconstruction technique for measuring the direct effect of quantization of CS measurements on the final sparse signal reconstruction quality, and (2) the analysis step decides appropriate quantized values to maximize the final sparse signal reconstruction quality. Through simulations, we compare the performance of the proposed quantization algorithms vis-à-vis existing quantization schemes.

I. INTRODUCTION

Based on a model of under-determined linear set of equations, compressed sensing (CS) [1] aims to reconstruct a high-dimensional sparse source vector, where most of coefficients are zero, from an under-sampled low-dimensional measurement vector. With a limited number of measurements (or a limited sampling resource), CS has emerged as a new powerful tool for sparse signal acquisition, compression and reconstruction. For many practical applications, CS measurements need to be quantized into a finite resolution representation, for subsequent transmission or storage. Later, using the finite resolution measurements, sparse signal reconstruction is performed, potentially followed by other inference tasks.

In this paper, we consider application scenarios where both measurement (or sampling) and quantization resources are limited. In particular, we assume that the total number of available bits for quantization is limited. One scenario would be in wireless body sensor networks (WBSN) [2] where a low number of measurements and a low bit-rate are available due to sensing costs and channel transmission rate restrictions, respectively. *Considering availability of limited number of measurements and quantization bits, we design new quantization algorithms with the objective of achieving high quality sparse signal reconstruction from the quantized CS measurements.*

A. Related Works

CS with quantized measurements has recently started to gain significant attention in the literature, and the available approaches are of three main kinds, described below.

- 1) In the first category, extensions to existing sparse reconstruction schemes have been studied without changing the quantization algorithms. Sparse signal reconstruction from noisy measurements – which can be thought of as the effect of quantization when the noise is bounded and additive – has been addressed in [3]. In [4], CS with finely quantized measurements using Sigma-Delta modulation is studied. In [5], the authors focus on convex optimization-based CS reconstruction from a set of quantized measurements, whereas [6] considers a greedy search algorithm for this purpose. In [7], a reconstruction scheme for more precise reconstruction of sparse signals from quantized measurements has been developed. In addition, robust schemes (against measurement noise) for reconstructing a sparse source from 1-bit quantized measurements have been proposed in [8], [9], and the design of message-passing algorithms for estimation from quantized CS samples has been studied in [10].
- 2) The second category considers the trade-off between the aspects of quantization (e.g., quantization bit-rate) and CS (e.g., number of measurements and loss in sparse reconstruction). In [11], high bit-rate theoretical bounds on average *end-to-end distortion* due to quantization of sparse sources and CS measurements have been derived, whereas the goal in [6] is to find high bit-rate average *quantization distortion* bounds caused by vector and scalar quantization of CS measurements. The aim in [12] is to analyze a trade-off between number of measurements and quantization bit-rate by introducing quantization compression regime versus CS compression regime.
- 3) Finally, in the third category, the main focus is on quantizer design for CS measurements while sparse reconstruction schemes are fixed. The design of high bit-rate quantizers for CS measurements that are optimal with respect to mean square error (MSE) of a particular convex optimization-based reconstruction method (LASSO [13]) is studied in [14]. Moreover, an optimal high rate quantizer design under message-passing reconstruction algorithms has been presented in [15]. In [16], a framework for scalar quantization of CS measurements has been proposed which provides exponential decay for instantaneous signal reconstruction distortion as a function of bit-rate.

B. Contributions of the Paper

The main contribution of this paper is in the third category mentioned above, i.e., quantizer design for CS measurements while sparse reconstruction algorithms are fixed. For this purpose, choosing the MSE as a performance criterion, we first derive necessary encoding conditions so as to minimize reconstruction MSE for a sparse input vector under a fixed decoder and a sparse reconstruction algorithm. Then, in order to feasibly implement the resulting rules, we develop a new framework for scalar quantization of CS measurements with the objective of achieving a lower *end-to-end reconstruction distortion* for the sparse source rather than *quantization distortion* for CS measurements. Technically, given a fixed quantizer look-up table and a fixed (but generic) sparse reconstruction scheme, the proposed algorithms strategically use a two-step mechanism in a closed-loop fashion: (1) the *synthesis* step employs a sparse signal reconstruction technique for measuring the direct effect of quantization of CS measurements on the final sparse signal reconstruction quality, and (2) the *analysis* step is performed followed by the synthesis step in order to choose appropriate quantized values to minimize the final sparse signal reconstruction distortion. This closed-loop strategy is known as *analysis-by-synthesis* (AbS) which has been widely used in multi-media coding [17]–[19].

To the best of our knowledge, the AbS approach has not been used for quantization of CS measurements earlier. The use of AbS principle is shown to provide a significantly better reconstruction performance compared to schemes that only consider direct quantization of CS measurements. Here, we mention that AbS requires higher computation. We analyze the computational complexity of the proposed algorithms, where it is shown that the complexity depends on availability of two compression resources, i.e., quantization bit-rate and number of CS measurements. As a byproduct, we also propose a low complexity scheme based on quantization of estimated sparsity patterns at the quantizer encoder which performs well at high quantization bit-rates. Further, we develop an *adaptive* quantization method by combining the proposed schemes in order to provide high-quality performance at all ranges of quantization and measurement rates. We experimentally evaluate the performance of our proposed algorithms, and also compare them with that of existing schemes for quantization of CS measurements.

C. Outline of the Paper

The remaining parts of the paper are organized as follows. In Section II, we give the problem statement which involves the CS and quantization models and also performance criterion. In Section III, we propose the new algorithms, and analyze their computational complexities. Thereafter, we show an adaptive quantization scheme in Section IV. The numerical results are given in Section V, and the conclusions are drawn in Section VI.

D. Notations

Scalar random variables (RV's) will be denoted by upper-case letters while their realizations (instants) will be de-

noted by the respective lower-case letters. Random vectors of dimension n will be represented by boldface characters. Hence, if \mathbf{Z} denotes a random row vector $[Z_1, \dots, Z_n]$, then, $\mathbf{z} = [z_1, \dots, z_n]$ indicates a specific realization of \mathbf{Z} . We will denote a sequence of RV's J_1, \dots, J_N by \mathbf{J}_1^N , further, $\mathbf{J}_1^N = \mathbf{j}_1^N$ implies that $J_1 = j_1, \dots, J_N = j_N$. Matrices will be denoted by capital Greek letters, and their pseudo-inverse by $(\cdot)^\dagger$. Further, a set is shown by a calligraphic character and its cardinality by $|\cdot|$, for example, $\mathcal{A} = \{a_i\}_{i=1}^n$ represents a set with cardinality $|\mathcal{A}| = n$. We will also denote the transpose of a vector by $(\cdot)^T$. We will use $\mathbb{E}[\cdot]$ to denote the expectation operator. The ℓ_p -norm ($p > 0$) of a vector \mathbf{z} will be denoted by $\|\mathbf{z}\|_p = (\sum_{n=1}^N |z_n|^p)^{1/p}$. $\|\mathbf{z}\|_0$ represents ℓ_0 -norm which is the number of non-zero coefficients in \mathbf{z} . Finally, for two real functions $f(n)$ and $g(n)$, $f(n) = \mathcal{O}(g(n))$ if $\exists c \in \mathbb{R}^+$ and $0 \leq f(n) \leq cg(n)$.

II. PROBLEM STATEMENT

In this section, we first introduce the preliminaries of CS and quantization, and then the design criterion and goal will be described.

A. Preliminaries on the CS Framework

The CS context can be thought of as a source compression method where a high-dimensional vector is mapped to a lower-dimensional vector both belonging to arbitrary continuous sets. Formally, we let a random sparse (in a fixed basis) signal $\mathbf{X} \in \mathbb{R}^M$ be linearly encoded using a known sensing matrix $\Phi \in \mathbb{R}^{N \times M}$ ($N < M$) representing a measurement (sampling) system which results in an under-determined set of linear measurements, i.e.,

$$\mathbf{Y} = \Phi \mathbf{X} \in \mathbb{R}^N. \quad (1)$$

We let \mathbf{X} be a K -sparse vector, i.e., it has at most K ($K < N$) non-zero coefficients, where the location of the non-zero's are uniformly drawn from all $\binom{M}{K}$ possibilities, and the magnitude of the non-zero coefficients are identically and independently drawn from a known distribution. We define the support set of the sparse vector $\mathbf{X} = [X_1, \dots, X_M]^T$ by $\mathcal{S} \triangleq \{m : X_m \neq 0\} \subset \{1, \dots, M\}$ and $|\mathcal{S}| = \|\mathbf{X}\|_0 \leq K$.

For the purpose of estimating a sparse vector from (possibly noisy) measurements, several efficient techniques have been developed based on convex optimization methods (see e.g. [3], [20], [21]), iterative greedy search algorithms (see e.g. [22]–[26]) and Bayesian estimation approaches (see e.g. [27]–[31]). For example, the goal of iterative greedy search algorithms is to first detect the sparsity pattern, i.e. location of non-zero coefficients, and then estimate the unknown non-zero coefficients. *In this paper, the design we present is generic in the sense that it works for any (fixed) sparse reconstruction algorithm.* We denote a sparse reconstruction algorithm by a mapping function $\mathbf{R} : \mathbb{R}^N \rightarrow \mathbb{R}^M$ which takes a measurement vector in N -dimensional space as an input, and produces an estimate of the sparse source vector in M -dimensional space ($N < M$) through a highly non-linear procedure. We also assume that the sensing matrix Φ is provided for the sparse reconstruction.

B. Quantization of CS Measurements

Scalar quantization is the process of mapping a sample from a continuous set to a discrete alphabet. We consider scalar quantization of the random CS measurements Y_n 's ($n = 1, \dots, N$). For this purpose, quantization is divided into *encoding* and *decoding* tasks. We consider a scalar *quantizer encoder* which maps each measurement to an appropriate index in a finite integer set in order for a *quantizer decoder* to make an estimate of the measurements based on the received index and a known decoding look-up table. We assume that the total bit budget (rate) allocated for quantization is $R_x \triangleq Mr_x$ bits per vector \mathbf{X} in which $r_x \in \mathbb{R}^+$ is the assigned quantization bit-rate to a scalar component of \mathbf{X} . Having the observations $\mathbf{Y} = \Phi\mathbf{X}$, each scalar Y_n ($n = 1, \dots, N$) of the measurement vector \mathbf{Y} is encoded by $r_y \triangleq Mr_x/N$ bits.¹ For each entry Y_n , the quantizer encoder is defined by a mapping $\mathbf{E} : \mathbb{R} \rightarrow \mathcal{I}$, where \mathcal{I} denotes the index set defined as $\mathcal{I} \triangleq \{0, 1, \dots, 2^{r_y} - 1\}$ with $|\mathcal{I}| = 2^{r_y}$. Denoting the quantized index by the RV I_n ($n = 1, \dots, N$), the encoder works according to $Y_n \in \mathcal{R}^{i_n} \Rightarrow I_n = i_n$, where the sets $\{\mathcal{R}^{i_n}\}_{i_n=0}^{2^{r_y}-1}$ are called *encoder regions* and $\bigcup_{i_n=0}^{2^{r_y}-1} \mathcal{R}^{i_n} = \mathbb{R}$. Next, we define the quantizer decoder which is characterized by a mapping $\mathbf{D} : \mathcal{I} \rightarrow \mathcal{C}_n$. The quantizer decoder takes the index I_n , and performs according to an available look-up table; $I_n = i_n \Rightarrow \hat{Y}_n = c_{i_n}$. Note that \hat{Y}_n is the quantized measurement RV associated with the entry Y_n , and the set of all reproduction *codepoints* $\mathcal{C}_n \triangleq \{c_{i_n}\}_{i_n=0}^{2^{r_y}-1}$ associated with the entry Y_n is called a *codebook*. Let us denote by $\hat{\mathbf{X}} \in \mathbb{R}^M$ the estimation of the input signal vector from quantized measurement vector $\hat{\mathbf{Y}} \triangleq [\hat{Y}_1, \dots, \hat{Y}_N]^T$ using a sparse reconstruction function \mathbf{R} . Then in a more compact way, given a fixed reconstruction \mathbf{R} , we state acquisition, quantized transmission and reconstruction equations as

$$\begin{aligned} \mathbf{Y} &= \Phi\mathbf{X}, \\ I_n &= \mathbf{E}(Y_n), \hat{Y}_n = \mathbf{D}(I_n) = c_{I_n}, \forall n \\ \hat{\mathbf{X}}(I_1, \dots, I_N) &\triangleq \hat{\mathbf{X}}(\mathbf{I}_1^N) = \mathbf{R}([c_{I_1}, \dots, c_{I_N}]^T). \end{aligned} \quad (2)$$

C. Objectives and Preliminary Analysis

We establish the goal of our work as follows.

Problem 1: For the system setup (2) with fixed codebook sets $\mathcal{C}_n = \{c_{i_n}\}_{i_n=0}^{2^{r_y}-1}$ ($n = 1, \dots, N$), a bit budget $R_x = Mr_x = Nr_y$, the objective is to implement the encoder, i.e., to find *encoding indexes* (transmission indexes) $i_n \in \mathcal{I}$ ($n = 1, \dots, N$) such that the *end-to-end* MSE, $\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2]$, is minimum. Therefore, we address the following optimization problem

$$\{i_1^*, \dots, i_N^*\} = \arg \min_{\{i_n \in \mathcal{I}\}_{n=1}^N} \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2], \quad (3)$$

where $\{i_n^*\}_{n=1}^N$ are the optimal encoding indexes (with respect to minimizing the end-to-end MSE) for quantization of the measurement vector $\mathbf{Y} = [Y_1, \dots, Y_N]^T$.

¹In practice, r_y can be a non-integer value, however, in the design procedure we let r_y be a positive integer, and later in simulation results, we show how to address the non-integer issue.

To clarify the objective, we first introduce a common encoding approach, referred to as *nearest-neighbor coding* for CS measurements. In this type of quantization, each scalar entry of the measurement vector is coded to its nearest codepoint. Therefore, given a fixed codebook associated with the measurement entry $Y_n = y_n$, and given that $\hat{Y}_n = c_{i_n}$, the nearest-neighbor quantizer/encoder uses the following encoding rule

$$i_n^* = \arg \min_{i_n \in \mathcal{I}} |y_n - c_{i_n}|^2, \forall n, \quad (4)$$

which minimizes the MSE per measurement entry, i.e., $\mathbb{E}[\|Y_n - \hat{Y}_n\|^2]$. However, this approach does not necessarily guarantee that the end-to-end MSE (the final performance measure) $\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2]$ is also minimized subject to fixed codebook sets. This is due to non-linear behavior of the sparse reconstruction function \mathbf{R} and non-orthogonality of the CS sensing matrix Φ . We also mention that, in this work, the decoder codebooks \mathcal{C}_n ($n = 1, \dots, N$) are designed off-line, and we do not address the separate issue of codebook design; the codebooks are given and fixed.

Now we show how MSE-minimizing encoding indexes are chosen. First, let us denote the *minimum mean square error* (MMSE) estimation of \mathbf{X} given the measurements $\mathbf{Y} = \mathbf{y}$ by

$$\tilde{\mathbf{x}}(\mathbf{y}) \triangleq \mathbb{E}[\mathbf{X}|\mathbf{Y} = \mathbf{y}] \in \mathbb{R}^M. \quad (5)$$

Next, using the notations introduced in Section I-D, we rewrite the end-to-end MSE, $\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2]$, as (6) on top of next page, where (a) and (b) are followed by marginalization over \mathbf{Y} and \mathbf{I}_1^N , respectively. Also, (c) follows from interchanging the integral and summation and the fact that $\Pr\{\mathbf{I}_1^N = \mathbf{i}_1^N | \mathbf{Y} = \mathbf{y}\} = 1, \forall y_1 \in \mathcal{R}^{i_1}, \dots, y_N \in \mathcal{R}^{i_N}$, and otherwise the probability is zero. Here, $f(\mathbf{y})$ is the N -fold probability density function (p.d.f.) of the measurement vector.

Then, using the fixed codebooks \mathcal{C}_n ($n = 1, \dots, N$), the MSE-minimizing encoding indexes are identical to finding the ones which minimize the term in the braces in the last expression of (6) since $f(\mathbf{y})$ is always non-negative. The resulting optimal indexes denoted by $\{i_n^* \in \mathcal{I}\}_{n=1}^N$ are obtained by (7), where (a) holds since the conditional expectation of sum of RV's equals to the sum of their conditional expectations. (b) follows from the fact that \mathbf{X} is independent of \mathbf{I}_1^N , conditioned on \mathbf{Y} , and hence, $\mathbb{E}[\|\mathbf{X}\|_2^2 | \mathbf{Y} = \mathbf{y}, \mathbf{I}_1^N = \mathbf{i}_1^N] = \mathbb{E}[\|\mathbf{X}\|_2^2 | \mathbf{Y} = \mathbf{y}]$ which is pulled out of the optimization. Also, (c) follows from a similar rationale, i.e., $\hat{\mathbf{X}}(\mathbf{I}_1^N)$ is independent of \mathbf{Y} , conditioned on \mathbf{I}_1^N . Further, \mathbf{X} and $\hat{\mathbf{X}}(\mathbf{I}_1^N)$ are independent conditioned on \mathbf{Y} and \mathbf{I}_1^N . For a realization $\mathbf{Y} = \mathbf{y}$, the last expression in (7) can be rewritten as (using (5))

$$\{i_1^*, \dots, i_N^*\} = \arg \min_{\mathbf{i}_1^N} \{ \|\hat{\mathbf{x}}(\mathbf{i}_1^N)\|_2^2 - 2\tilde{\mathbf{x}}(\mathbf{y})^T \hat{\mathbf{x}}(\mathbf{i}_1^N) \}, \quad (8)$$

where \mathbf{i}_1^N denotes the sequence i_1, \dots, i_N .

Unfortunately, solving (8) jointly for all encoding indexes is not analytically and practically tractable for a general sparse reconstruction function as it requires reconstruction and searching over all possible codepoints, leading to prohibitive complexity. Since we are interested in developing a reasonably simple coding system, we refrain from the

$$\begin{aligned}
\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2] &= \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}(\mathbf{I}_1^N)\|_2^2] \stackrel{(a)}{=} \int_{\mathbf{y}} \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}(\mathbf{I}_1^N)\|_2^2 | \mathbf{Y} = \mathbf{y}] f(\mathbf{y}) d\mathbf{y} \\
&\stackrel{(b)}{=} \int_{\mathbf{y}} \sum_{i_1} \dots \sum_{i_N} \Pr\{\mathbf{I}_1^N = \mathbf{i}_1^N | \mathbf{Y} = \mathbf{y}\} \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}(\mathbf{I}_1^N)\|_2^2 | \mathbf{Y} = \mathbf{y}, \mathbf{I}_1^N = \mathbf{i}_1^N] f(\mathbf{y}) d\mathbf{y} \\
&\stackrel{(c)}{=} \sum_{i_1} \dots \sum_{i_N} \int_{\mathbf{y}_1 \in \mathcal{R}^{i_1}} \dots \int_{\mathbf{y}_N \in \mathcal{R}^{i_N}} \left\{ \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}(\mathbf{I}_1^N)\|_2^2 | \mathbf{Y} = \mathbf{y}, \mathbf{I}_1^N = \mathbf{i}_1^N] \right\} f(\mathbf{y}) d\mathbf{y}
\end{aligned} \tag{6}$$

$$\begin{aligned}
\{i_n^*\}_{n=1}^N &= \arg \min_{\mathbf{i}_1^N} \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}(\mathbf{I}_1^N)\|_2^2 | \mathbf{Y} = \mathbf{y}, \mathbf{I}_1^N = \mathbf{i}_1^N] \\
&\stackrel{(a)}{=} \arg \min_{\mathbf{i}_1^N} \left\{ \mathbb{E}[\|\mathbf{X}\|_2^2 | \mathbf{Y} = \mathbf{y}, \mathbf{I}_1^N = \mathbf{i}_1^N] + \mathbb{E}[\|\hat{\mathbf{X}}(\mathbf{I}_1^N)\|_2^2 | \mathbf{Y} = \mathbf{y}, \mathbf{I}_1^N = \mathbf{i}_1^N] - 2\mathbb{E}[\mathbf{X}^T \hat{\mathbf{X}}(\mathbf{I}_1^N) | \mathbf{Y} = \mathbf{y}, \mathbf{I}_1^N = \mathbf{i}_1^N] \right\} \\
&\stackrel{(b)}{=} \arg \min_{\mathbf{i}_1^N} \left\{ \mathbb{E}[\|\hat{\mathbf{X}}(\mathbf{I}_1^N)\|_2^2 | \mathbf{Y} = \mathbf{y}, \mathbf{I}_1^N = \mathbf{i}_1^N] - 2\mathbb{E}[\mathbf{X}^T \hat{\mathbf{X}}(\mathbf{I}_1^N) | \mathbf{Y} = \mathbf{y}, \mathbf{I}_1^N = \mathbf{i}_1^N] \right\} \\
&\stackrel{(c)}{=} \arg \min_{\mathbf{i}_1^N} \left\{ \mathbb{E}[\|\hat{\mathbf{X}}(\mathbf{I}_1^N)\|_2^2 | \mathbf{I}_1^N = \mathbf{i}_1^N] - 2\mathbb{E}[\mathbf{X}^T | \mathbf{Y} = \mathbf{y}] \mathbb{E}[\hat{\mathbf{X}}(\mathbf{I}_1^N) | \mathbf{I}_1^N = \mathbf{i}_1^N] \right\}
\end{aligned} \tag{7}$$

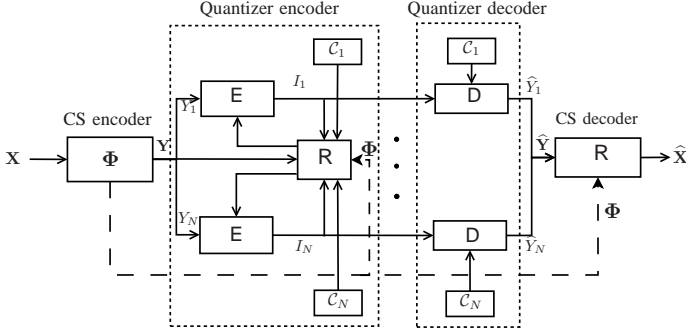


Fig. 1. Analysis-by-synthesis (AbS) quantization of CS measurements. At the quantizer encoder, the function \mathbf{R} performs as a *synthesizer* based on a sparse reconstruction, and the function \mathbf{E} acts as an *analyzer* based on the optimization problem (9) in a closed-loop.

optimal joint encoder. Instead, in this work, we focus on sub-optimal techniques (with respect to (8)) for quantization of CS measurements. The quantization schemes are developed in the next section.

III. ANALYSIS-BY-SYNTHESIS QUANTIZATION OF CS MEASUREMENTS

In this section, we first show how an encoding index can be chosen by fixing the other indexes under the assumptions of Problem 1, and then develop AbS-based quantization schemes following by complexity analysis. Our studied AbS system is illustrated in Figure 1.

In order to solve (8) approximately, we consider optimizing one variable by fixing the others, that is, optimizing the index i_n by fixing the indexes $i_1, \dots, i_{n-1}, i_{n+1}, \dots, i_N$. This is indeed an *alternating* optimization approach and sub-optimal compared to the joint optimization method (8), but provides a feasible solution with low complexity.

Now, assume that given the codebook sets, we fix all encoding indexes except i_n ($n \in \{1, \dots, N\}$). Then, from

(8), the encoder attains the following encoding rule

$$i_n^* = \arg \min_{i_n \in \mathcal{I}} \left\{ \|\hat{\mathbf{x}}(i_n)\|_2^2 - 2\tilde{\mathbf{x}}(\mathbf{y})^T \hat{\mathbf{x}}(i_n) \right\}. \tag{9}$$

By denoting $\hat{\mathbf{x}}(i_n) \triangleq \mathbf{R}([c_{i_1}, \dots, c_{i_n}, \dots, c_{i_N}])^T$, we refer that the reconstructed signal is dependent only upon the index (codepoints) associated with the n^{th} measurement entry. Interestingly, (9) indicates that in order to minimize the end-to-end MSE for a fixed codebook, the index i_n should be chosen such that the final reconstruction vector is as close as possible (in ℓ_2 -norm) to the MMSE estimation of the sparse source given the measurements. We employ (9) in an iterate-alternate optimization approach to realize new AbS-based algorithms which are described in the next subsection. Before going to the next subsection, we have the following remark.

Remark 1: As shown in our formulations (e.g. (9)), we need the MMSE estimate $\tilde{\mathbf{x}}(\mathbf{y})$ to find the encoding indexes. However, in practice, implementing the MMSE estimator is not feasible. Therefore, we approximate the MMSE estimate $\tilde{\mathbf{x}}(\mathbf{y})$ by the sparse signal estimate of reconstruction function \mathbf{R} using unquantized measurement \mathbf{y} . Denoting the sparse signal estimate by $\tilde{\mathbf{x}}(\mathbf{y})$, we assume $\tilde{\mathbf{x}}(\mathbf{y}) \approx \hat{\mathbf{x}}(\mathbf{y})$. In this paper, we use the greedy *orthogonal matching pursuit* (OMP) [22], [23] reconstruction algorithm as the sparse reconstruction function \mathbf{R} . The choice of OMP is motivated due to a good tradeoff between complexity and reconstruction performance. The OMP algorithm is briefly described in the Appendix. We emphasize that our formulation is general and does not deter use of any practical reconstruction algorithm in lieu of OMP.

A. Proposed Quantization Algorithms

We first describe the *iterative* framework for the proposed quantization schemes summarized in Algorithm 1. Suppose that the codebook sets \mathcal{C}_n ($n = 1, \dots, N$) are designed offline, and let the quantizer encoder have access to the sensing matrix

Φ as well as the codebooks (step (1)).¹ In step (2), we obtain the locally reconstructed vector $\bar{\mathbf{x}}(\mathbf{y})$ as an approximation to the MMSE estimator $\hat{\mathbf{x}}(\mathbf{y})$. Now, we define a dummy vector $\mathbf{z} \in \mathbb{R}^N$, where its n^{th} component is chosen uniformly at random from the set \mathcal{C}_n ($\forall n$) at the first iteration (step (3)). The vector \mathbf{z} stores coefficients of quantized CS measurements. Throughout iterations, the entries of \mathbf{z} are either *sequentially* or *non-sequentially* adjusted for minimizing the reconstruction MSE. Now, we describe the subroutine $\text{AbS}(\cdot)$ executed in Algorithm 1 (step (6)) which uses two different *alternating* approaches (in terms of performance and complexity) for implementing (9).

Algorithm 1 : AbS quantization

- 1: **input:** $\mathcal{C}_n = \{c_{i_n}\}_{i_n=0}^{2^{r_y}-1}$ ($\forall n = 1, \dots, N$) and Φ, \mathbf{y}, γ (stopping threshold)
 - 2: **compute:** $\bar{\mathbf{x}}(\mathbf{y})$ (locally reconstructed vector)
 - 3: **initialize** $\mathbf{z}^{(0)} \in \mathbb{R}^N$, where $z_n^{(0)} \in \mathcal{C}_n, \forall n$.
 - 4: Set $l \leftarrow 0$ (iteration counter)
 - 5: **repeat**
 - 6: $[i_n^*, \hat{\mathbf{x}}^{(l+1)}(i_n^*), \mathbf{z}^{(l+1)}] = \text{AbS}(\mathcal{C}_n, \bar{\mathbf{x}}(\mathbf{y}), \mathbf{z}^{(l)}), \forall n$
 - 7: $l \leftarrow l + 1$
 - 8: **until** $|\|\hat{\mathbf{x}}^{(l)}(i_n^*)\|_2^2 - 2\bar{\mathbf{x}}(\mathbf{y})^T \hat{\mathbf{x}}^{(l)}(i_n^*)|$
 $- |\|\hat{\mathbf{x}}^{(l-1)}(i_n^*)\|_2^2 + 2\bar{\mathbf{x}}(\mathbf{y})^T \hat{\mathbf{x}}^{(l-1)}(i_n^*)| < \gamma, \forall n$
 - 9: **output:** $I_n = i_n^*, \hat{Y}_n = c_{i_n^*}, \forall n$
-

- 1) **Sequential AbS quantization:** Our first AbS-based quantization scheme which can be executed in Algorithm 1 is summarized in the Subroutine $\text{AbS_seq}(\cdot)$ where the main idea is that each measurement entry is sequentially adjusted towards the direction of its MSE-minimizing codepoint at each iteration.

Subroutine: $\text{AbS_seq}(\mathcal{C}_n, \bar{\mathbf{x}}(\mathbf{y}), \mathbf{z}^{(l)})$

- 1: **for** $n = 1 : N$ **do**
 - 2: **for** $i = 0 : 2^{r_y} - 1$ **do**
 - 3: $z_n^{(l)} \leftarrow c_{i_n}$
 - 4: **compute:** $\hat{\mathbf{x}}^{(l)}(i_n) = \mathbf{R}(\mathbf{z}^{(l)})$
 - 5: **end for**
 - 6: $i_n^* = \arg \min_{i_n \in \mathcal{I}} \{\|\hat{\mathbf{x}}^{(l)}(i_n)\|_2^2 - 2\bar{\mathbf{x}}(\mathbf{y})^T \hat{\mathbf{x}}^{(l)}(i_n)\}$
 - 7: **update:** $z_n^{(l)} \leftarrow c_{i_n^*}(i_n)$
 - 8: **end for**
 - 9: **output:** $i_n^*, \hat{\mathbf{x}}^{(l)}(i_n^*), \mathbf{z}^{(l)}$
-

Using Algorithm 1, the function $\text{AbS_seq}(\cdot)$ accepts the codebooks $\mathcal{C}_n, \forall n$, the locally reconstructed vector $\bar{\mathbf{x}}(\mathbf{y})$ and the dummy vector $\mathbf{z}^{(l)}$. At iteration l , the n^{th} ($n = 1, \dots, N$) entry of $\mathbf{z}^{(l)}$, denoted by $z_n^{(l)}$, is

¹Note that the sparsity level K may be also provided at the encoder and the decoder if the OMP algorithm is used as the reconstruction function \mathbf{R} . If subspace pursuit [24] or CoSaMP [25] reconstruction algorithms are used, then K must be provided. However, using the basis pursuit [20] or LASSO [13] reconstruction algorithms, the sparsity level is not necessarily required.

replaced by all 2^{r_y} codepoints from the set \mathcal{C}_n (step (3)) while the other entries are fixed, and the reconstructed vectors, denoted by $\hat{\mathbf{x}}^{(l)}(i_n) = \mathbf{R}(\mathbf{z}^{(l)})$ ($i_n \in \mathcal{I} = \{0, \dots, 2^{r_y} - 1\}$), are synthesized corresponding to each vector (step (4)). Then, an optimization is carried out by solving $\arg \min_{i_n \in \mathcal{I}} \{\|\hat{\mathbf{x}}^{(l)}(i_n)\|_2^2 - 2\bar{\mathbf{x}}(\mathbf{y})^T \hat{\mathbf{x}}^{(l)}(i_n)\}$ so as to

find the winning MSE-minimizing encoding index i_n^* (step (6)). Next, the n^{th} entry of the vector $\mathbf{z}^{(l)}$ is updated by the codepoint associated with the analyzed index, i.e., $c_{i_n^*}$ (step (7)). This procedure continues for each entry of $\mathbf{z}^{(l)}$ sequentially, and the subroutine produces the optimized encoding index i_n^* , and the reconstructed vector $\hat{\mathbf{x}}^{(l)}(i_n^*)$ as well as the updated quantized vector $\mathbf{z}^{(l)}$ (step (9)) which will be used by the subroutine at the next iteration of Algorithm 1.

- 2) **Non-sequential AbS quantization:** Using Algorithm 1, our second proposed Subroutine $\text{AbS_nonseq}(\cdot)$ accepts the codebooks $\mathcal{C}_n, \forall n$, the locally reconstructed vector $\bar{\mathbf{x}}(\mathbf{y})$ and the dummy vector $\mathbf{z}^{(l)}$. This non-sequential scheme is not order-dependent, and finds the MSE-minimizing index/codepoint by tracking the best path at each iteration.

Subroutine: $\text{AbS_nonseq}(\mathcal{C}_n, \bar{\mathbf{x}}(\mathbf{y}), \mathbf{z}^{(l)})$

- 1: $\mathcal{L} = \emptyset$
 - 2: **repeat**
 - 3: **initialize:** $\Delta = \mathbf{0}^{(N-|\mathcal{L}|) \times 2^{r_y}}$
 - 4: **for** $n \in \{1, \dots, N\} \setminus \mathcal{L}$ **do**
 - 5: **for** $i = 0 : 2^{r_y} - 1$ **do**
 - 6: $z_n^{(l)} \leftarrow c_{i_n}$
 - 7: **compute:** $\hat{\mathbf{x}}^{(l)}(i_n) = \mathbf{R}(\mathbf{z}^{(l)})$
 - 8: **end for**
 - 9: **compute:** $\Delta(n, i + 1) = \|\hat{\mathbf{x}}^{(l)}(i_n)\|_2^2 - 2\bar{\mathbf{x}}(\mathbf{y})^T \hat{\mathbf{x}}^{(l)}(i_n)$
 - 10: **end for**
 - 11: $[n^*, i_n^*] = \arg \min_{n, i} \Delta(n, i + 1)$
 - 12: **update:** $z_{n^*}^{(l)} \leftarrow c_{i_n^*}$
 - 13: $\mathcal{L} \leftarrow \mathcal{L} \cup \{n^*\}$
 - 14: **until** $\mathcal{L} = \{1, 2, \dots, N\}$
 - 15: **output:** $i_n^*, \hat{\mathbf{x}}^{(l)}(i_n^*), \mathbf{z}^{(l)}$
-

At the first iteration of $\text{AbS_nonseq}(\cdot)$, we define a set \mathcal{L} which is initially empty (step (1)). At iteration l of the algorithm, the n^{th} ($n = 1, \dots, N$) entry of $\mathbf{z}^{(l)}$, denoted by $z_n^{(l)}$, is replaced by all 2^{r_y} codepoints in the set \mathcal{C}_n (step (6)) while the other entries are fixed, and the reconstructed vectors, denoted by $\hat{\mathbf{x}}^{(l)}(i_n) = \mathbf{R}(\mathbf{z}^{(l)})$ ($i_n \in \mathcal{I} = \{0, \dots, 2^{r_y} - 1\}$), are synthesized corresponding to each vector (step (7)) whose values are stored in the matrix $\Delta \in \mathbb{R}^{(N-|\mathcal{L}|) \times 2^{r_y}}$ (step (9)). Then, an analysis is performed by solving (9), i.e., $\arg \min_{n, i} \{\|\hat{\mathbf{x}}^{(l)}(i_n)\|_2^2 - 2\bar{\mathbf{x}}(\mathbf{y})^T \hat{\mathbf{x}}^{(l)}(i_n)\}$, through a search in rows and columns of the matrix Δ , so as to find the MSE-minimizing

encoding index i_n^* and the entry's index n^* (step (11)). Now, the entry n^* of the vector $\mathbf{z}^{(l)}$ is updated by the codepoint associated with the analyzed index, i.e., $c_{i_n^*}$ (step (12)). Note that the set \mathcal{L} expands by adding n^* to the previous set (step (13)) in order to exclude the minimizing index and entry at the next iteration of the subroutine. The iterations continue until the set \mathcal{L} consists of all elements of the set $\{1, 2, \dots, N\}$ so that all entries are assigned by the minimizing codepoints (step (14)). At the last iteration of the subroutine, it outputs the optimized transmission index i_n^* , and the reconstructed vector $\hat{\mathbf{x}}^{(l)}(i_n^*)$ as well as the updated quantized vector $\mathbf{z}^{(l)}$ (step (15)) which will be used by the subroutine at the next iteration of Algorithm 1.

Algorithm 1 iterates until convergence where the stopping criterion is that reconstruction improvement at two consecutive iterations is smaller than a predefined threshold $\gamma > 0$. After convergence, the algorithm outputs the quantization indexes I_n 's and the CS quantized measurements \hat{Y}_n 's, $n = 1, \dots, N$, (step (9)) in which the latter is regarded as an input to the final sparse reconstruction algorithm in order to provide the estimate $\hat{\mathbf{x}}$. In the following, we discuss the convergence of Algorithm 1.

Remark 2: Following standard convergence proofs in [32, Lemmas 11.3.1-2], we provide qualitative arguments on the convergence of Algorithm 1. By construction (and ignoring issues such as numerical precision), the iterative design in Algorithm 1 given codebook sets \mathcal{C}_n ($n = 1, \dots, N$) and a fixed sparse reconstruction algorithm \mathbf{R} always converges to a local optimum. More precisely, at each iteration of Algorithm 1, given the fixed codebook sets and a sparse reconstruction algorithm, whenever the criteria in step (6) of `AbS_seq` or step (11) of `AbS_nonseq` are invoked, the reconstruction distortion given the updated index and the remaining fixed indexes can only leave unchanged or reduced. This is due to the fact that the distortion-minimizing index is always chosen. Hence, the distortion monotonically decreases at each iteration. Since the stopping criterion is defined as the difference in distortion at successive iterations, the stopping condition is bound to satisfy after finite number of iterations, and the algorithm converges to optimized encoding indexes. However, nothing can be generally guaranteed about global optimality.

B. Complexity Analysis

In this section, we analyze the encoding computational complexity of the proposed quantization schemes. We mainly quantify how many times a sparse reconstruction algorithm is invoked throughout the AbS procedures. Note that, in practice, other operators used in the proposed algorithms have negligible complexity compared to the sparse reconstruction algorithms.

First, recall from (8) that an exhaustive search for the joint optimization requires $\mathcal{O}(2^{Mr_x})$, or $\mathcal{O}(2^{Nr_y})$ (since $Mr_x = Nr_y$), computations of a sparse reconstruction algorithm. This is not feasible in practice. Employing the sequential AbS

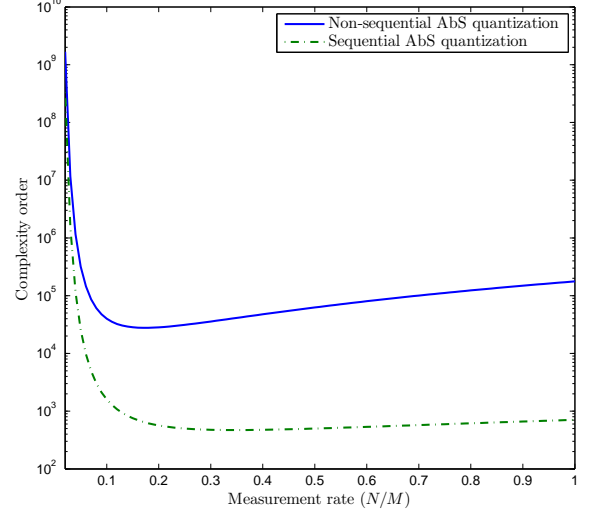


Fig. 2. Complexity order as a function of measurement rate $\frac{N}{M}$ for the proposed AbS quantization schemes for $M = 500$ and $r_x = 0.5$ bit/component.

quantization (Subroutine `AbS_seq`), the operations for calculating the encoding indexes increase at most like $\mathcal{O}(N2^{\frac{Mr_x}{N}})$ at each iteration of Algorithm 1, where it follows that at L iterations, the total complexity increases with $\mathcal{O}(LN2^{\frac{Mr_x}{N}})$. Next, let us consider the non-sequential AbS-based quantization (Subroutine `AbS_nonseq`) at one iteration of Algorithm 1. The operations for calculating encoding indexes require $N2^{Mr_x/N} + (N-1)2^{Mr_x/N} + \dots + 2^{Mr_x/N} = \frac{N(N+1)}{2}2^{Mr_x/N}$ computations of a sparse reconstruction algorithm since at each iteration of the subroutine the set \mathcal{L} expands by adding one element. Hence, it can be shown that the computational complexity of the non-sequential AbS quantization after L iterations of Algorithm 1 grows at most like $\mathcal{O}(L\frac{N^2}{2}2^{\frac{Mr_x}{N}})$.

The order of computations for the AbS schemes indicates a tradeoff between availability of compression resources, i.e., number of measurements and quantization bit budget. It also implies that, for a fixed bit budget $R_x = Mr_x$, by increasing the number of measurements, first the complexity decays sharply, and then at some point it starts increasing with a small slope. This is due to the fact that the complexity depends on the number of measurements (through the linear term N or the quadratic term N^2 for the sequential and the non-sequential AbS-based algorithms, respectively) and the quantization bit-rate per measurement entry $r_y = Mr_x/N$ (through the exponential term $2^{\frac{Mr_x}{N}}$). This aspect is shown in Figure 2.

We finalize this section with a remark regarding codebook design for the studied quantization schemes.

Remark 3: As mentioned in the design of the proposed schemes, we have assumed that the codebook sets are given. However, using unoptimized codebooks may lead to poor performance. Here, we describe an alternative to training codebooks corresponding to each quantization scheme. A standard

approach would be the *Lloyd algorithm* [32, Chapter 6], where, with possibly random initializations, either the encoder (index allocation) or decoder (codepoints) is assumed known and the other is selected optimally with respect to minimizing a particular distortion measure and a statistically specified input. This procedure is then alternated and iterated until (local) convergence is reached.

Due to the discrete mixture distribution of sparse source and highly non-linear behavior of sparse reconstruction algorithms, it is generally very challenging to derive closed-form optimal codepoints with respect to minimizing the end-to-end MSE. Therefore, we devise a potentially sub-optimal design for codebook training. Let us assume that CS measurements are identically and independently distributed (i.i.d.), and introduce the average quantization distortion $\mathbb{E}[|Y_n - \hat{Y}_n|^2]$ as a criterion. Then, all the codebooks are the same, and the codepoints are optimized using the Lloyd algorithm for a random measurement entry Y_n .¹ The fact that all codebooks are chosen the same for each measurement entry can be justified by similar observations in [33]. The optimized codepoints using the Lloyd algorithm would minimize the quantization distortion per measurement entry. Therefore, these codepoints can be considered *good* alternatives for the AbS-based quantizers which take the end-to-end distortion into account. The codebook sets are designed offline, and will be used for implementation of the quantization schemes whose performances are evaluated in Section V.

In what follows, we investigate coding of the signal domain where its combination with coding of the CS measurement domain leads to an adaptive design algorithm.

IV. ADAPTIVE CODING: SIGNAL AND MEASUREMENT DOMAINS

Up to this point, we have developed new schemes for quantization in the CS measurement domain. Since using the proposed AbS designs, we reconstruct the source vector according to which the quantization is performed, it is important to investigate a scenario where the locally reconstructed source $\bar{\mathbf{x}}$ (we drop the dependency of the vector on \mathbf{y} for simplicity of notation) is coded.

A. Signal Domain Coding

In order to efficiently address the design problem which aims to quantize in the domain of reconstructed signal, two alternatives may be visualized. They are as follows.

- 1) **Direct coding:** Using this scheme, all coefficients of the vector $\bar{\mathbf{x}} \in \mathbb{R}^M$ are quantized with available r_x bits per component. Therefore, in order to encode all coefficients, this scheme requires $R_x \geq M$ (or, $r_x \geq 1$) bits. Formally, let the codebook sets $\mathcal{G}_m \triangleq \{g_{i_m} \in \mathbb{R}\}_{i_m=0}^{2^{r_x}-1}$ corresponding to components \bar{x}_m 's ($m = 1, \dots, M$)

¹Note that in the case of exactly K -sparse vector whose non-zero components are i.i.d. RV's, using the central limit theorem, each entry of the linear measurement vector converges weakly to a Gaussian random variable with zero mean and variance K/N as $(K, M, N) \rightarrow \infty$ and with the rates K/M and N/M remain constant (see [6] for more details).

be given. Then, the encoding index associated with the component \bar{x}_m is chosen as

$$i_m^* = \arg \min_{i_m \in \mathcal{I}} |\bar{x}_m - g_{i_m}|^2, \quad m = 1, \dots, M, \quad (10)$$

where the index set \mathcal{I} is defined as $\mathcal{I} \triangleq \{0, \dots, 2^{r_x} - 1\}$. The decoder functions according to a look-up table $I_m = i_m \Rightarrow \hat{X}_m = g_{i_m}$. Using the direct coding scheme, all the codebooks are assumed the same (since the components of \mathbf{X} are i.i.d. RV's), and codepoints are optimized using the Lloyd algorithm by applying the performance measure $\mathbb{E}[|X_m - \hat{X}_m|^2]$.

- 2) **Support set coding:** In this case, we take into account the sparsity pattern of the locally reconstructed vector $\bar{\mathbf{x}}$. We first code the reconstructed support set, and then the magnitude of the non-zero coefficients on the support set. We denote by $\bar{\mathbf{x}}_{\hat{\mathcal{S}}} \in \mathbb{R}^K$ the entries of $\bar{\mathbf{x}}$ indexed by the elements of the estimated support set, denoted by $\hat{\mathcal{S}} \subset \{1, \dots, M\}$. Therefore, each component of $\hat{\mathcal{S}}$ can be represented by $\log_2 M$ bits that can be coded, and then recovered without loss. Now, we map the magnitude of K largest non-zero coefficients of $\bar{\mathbf{x}}$ to their nearest codepoints using $R_0 \triangleq R_x - K \log_2 M$ bits, where $R_x = M r_x$ is the total bit budget. Suppose the codebooks corresponding to the estimated non-zero coefficients, denoted by $\bar{x}_s, s \in \hat{\mathcal{S}}$, be given by $\mathcal{D} = \{d_{i_s} \in \mathbb{R}\}_{i_s=0}^{2^{R_0/K}-1}$. Then, the encoding index associated with entry \bar{x}_s is chosen as

$$i_s^* = \arg \min_{i_s \in \mathcal{K}} |\bar{x}_s - d_{i_s}|^2, \quad s \in \hat{\mathcal{S}}, \quad (11)$$

where the index set \mathcal{K} is defined as $\mathcal{K} \triangleq \{0, \dots, 2^{R_0/K} - 1\}$. The decoder works according to a look-up table $I_s = i_s \Rightarrow \hat{X}_s = d_{i_s}$. For quantizing all non-zero coefficients, this approach requires at least $R_x \geq K \log_2 M + K$ bits. Looking at the direct and support set coding, it can be inferred that these approaches require rather high bit budget for quantization. Exploiting the support set coding approach, the Lloyd algorithm is used for codebook training by adopting the performance criterion $\mathbb{E}[|X_s - \hat{X}_s|^2]$ for the input RV $X_s, s \in \mathcal{S}$, where X_s is a non-zero coefficient of \mathbf{X} drawn according to a known distribution.

Next, we show how to combine the signal domain coding schemes with the CS measurement domain quantization schemes in order to adaptively gain a better performance.

B. Adaptive Coding

Till now, we have proposed quantization schemes in CS measurement domain and signal domain. An engineering approach is to choose adaptively a better quantized signal (when it is compared with the locally reconstructed signal $\bar{\mathbf{x}}$) between the signal domain coding and the CS measurement domain coding schemes. This can be performed by assigning 2 flag bits for representing the proposed four schemes at the decoder, i.e.,

- 1) nearest-neighbor coding (CS measurement domain),
- 2) non-sequential AbS quantization (CS measurement domain),

- 3) direct coding (signal domain), and
- 4) support set coding (signal domain).

Then, the remaining $Mr_x - 2$ bits are used for quantization of each individual scheme. Since the computational complexity of the direct, support set and nearest-neighbor coding schemes are negligible compared to the non-sequential AbS quantization scheme, the total complexity of the adaptive coding grows at most as that of the non-sequential AbS quantization.

V. EXPERIMENTS AND RESULTS

In this section, we first demonstrate the performance of the proposed quantization schemes:

- 1) Nearest-neighbor coding (for quantizing in the CS measurement domain),
- 2) sequential AbS quantization (for quantizing in the CS measurement domain),
- 3) non-sequential AbS quantization (for quantizing in the CS measurement domain),
- 4) direct coding (for quantizing in the signal domain),
- 5) support set coding (for quantizing in the signal domain), and
- 6) adaptive coding (for quantizing in both domains).

Finally, we compare the performance of AbS-based quantizer vis-a-vis existing methods in the literature, specifically, the methods from [14].¹

We mainly quantify the MSE obtained by these schemes in terms of availability of compression resources, i.e., number of measurements and quantization bit-rate. Before showing simulation results, we state experimental setups in the next subsection.

A. Experimental Setups

We quantify the performance using normalized MSE (NMSE) defined as

$$\text{NMSE} \triangleq \frac{\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2]}{\mathbb{E}[\|\mathbf{X}\|_2^2]}. \quad (12)$$

In principle, the numerator of NMSE in (12) is computed by sample averaging over generated realizations of \mathbf{X} using Monte-Carlo simulations, and the denominator can be calculated exactly under the assumptions in our simulation setup. This calculation will be given in details later.

In addition, in order to measure the level of under-sampling, we define measurement rate ($0 < \alpha \leq 1$) as

$$\alpha \triangleq N/M. \quad (13)$$

Our simulation setup includes the following steps:

- 1) For given values of sparsity level K (assumed known in advance) and input vector size M , choose α , and round the number of measurements N to its nearest integer.
- 2) Randomly generate a set of exactly K -sparse vector \mathbf{X} where the support set \mathcal{S} with $|\mathcal{S}| = K$ is chosen uniformly at random over the set $\{1, 2, \dots, M\}$. The

¹In the spirit of reproducible results, we provide MATLAB codes for simulation of the AbS-based quantizers in the following website: www.ee.kth.se/~amishi/reproducible_research.html.

non-zero coefficients of \mathbf{X} are i.i.d. R.V.'s drawn from a known distribution. In our simulations, we use two mostly-common distributions for the non-zero coefficients: Gaussian and uniform. Based on the uniform sparsity pattern assumption, $\mathbb{E}[\|\mathbf{X}\|_2^2]$ (the denominator in (12)) can be analytically derived. It follows that

$$\begin{aligned} \mathbb{E}[\|\mathbf{X}\|_2^2] &= \sum_{m=1}^M \mathbb{E}[X_m^2] \\ &= \sum_{m=1}^M \Pr(m \in \mathcal{S}) \mathbb{E}[X_m^2 | m \in \mathcal{S}] + \Pr(m \notin \mathcal{S}) \mathbb{E}[X_m^2 | m \notin \mathcal{S}] \\ &\stackrel{(a)}{=} \sum_{m=1}^M \frac{K}{M} \mathbb{E}[X_m^2 | m \in \mathcal{S}] \end{aligned} \quad (14)$$

where (a) follows from the assumption of the uniformly distributed sparsity pattern, and also from the fact the second moments of the coefficients of \mathbf{X} that are not within the support set \mathcal{S} are zero. Now, note that $\mathbb{E}[X_m^2 | m \in \mathcal{S}]$ shows the second moment of X_m within the support set \mathcal{S} . Therefore, it can be easily shown that

- if the non-zero coefficients of \mathbf{X} are drawn from Gaussian distribution $\mathcal{N}(0, 1)$, then $\mathbb{E}[\|\mathbf{X}\|_2^2] = K$,
 - if the non-zero coefficients of \mathbf{X} are drawn from the uniform distribution $\mathcal{U}[-1, 1]$, then $\mathbb{E}[\|\mathbf{X}\|_2^2] = \frac{K}{3}$.
- 3) Randomly generate a set of sensing matrix Φ . We let the elements of the sensing matrix be $\Phi_{ij} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1/N)$, and then normalize the columns of Φ to unit-norm. Note that once Φ is generated, it remains fixed and known to the sparse reconstruction algorithm.
 - 4) Compute linear measurements $\mathbf{Y} = \Phi\mathbf{X}$ for each sparse data, and apply a sparse reconstruction algorithm (here OMP) to acquire the locally reconstructed vector $\hat{\mathbf{x}}$ as discussed in Remark 1.
 - 5) Choose the total quantization bit-rate $R_x = Mr_x$ bits/vector where $r_x \in \mathbb{R}^+$ is the allocated rate to a component of \mathbf{X} , and design codebook sets for each quantization scheme as described in Remark 3. In order to make a fair comparison and guarantee that the performance of the AbS-based quantization does not decline as compared with that of the nearest-neighbor coding, Algorithm 1 is initialized with the nearest-neighbor codepoints.
 - 6) Next, apply the quantization algorithms on the generated data \mathbf{X}, \mathbf{Y} , and assess NMSE by averaging over all data.

B. Experimental Results: Evaluation of the Proposed Schemes

In our simulations, for implementing Algorithm 1 using the subroutines `AbS_seq` and `AbS_nonseq`, we choose the stopping threshold $\gamma = 10^{-6}$, where we have observed that Algorithm 1 converges in at most 5 iterations for all simulation setups. We perform the simulation by averaging over 1000 realizations of \mathbf{X} with non-zero coefficients drawn from the standard Gaussian distribution. Furthermore, the OMP reconstruction algorithm is used to recover the source from quantized measurements.

Next, recall that the number of quantization bits assigned for each entry of $\mathbf{Y} \in \mathbb{R}^N$ is $r_y = \frac{Mr_x}{N}$, and may not

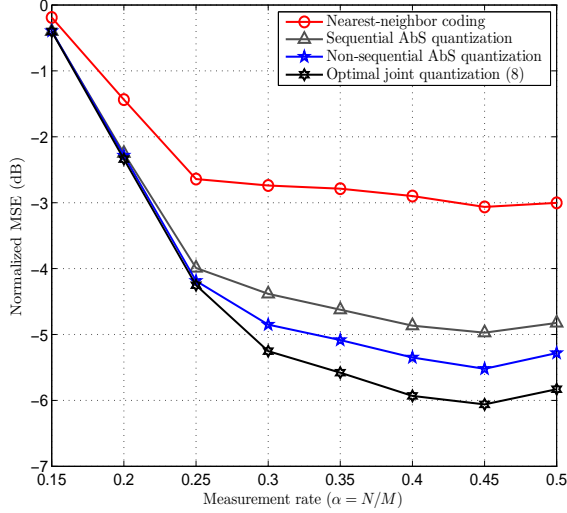


Fig. 3. Normalized MSE (in dB) as a function of measurement rate ($\alpha = N/M$) given a fixed quantization bit-rate using nearest-neighbor coding, sequential and non-sequential AbS quantization schemes and joint quantization (8) as a benchmark. The parameters are chosen as $M = 40$, $K = 3$ and $r_x = 0.5$ bit per component of the source vector.

be an integer. Hence, in order to utilize all available bits for quantization of CS measurements, we first assign $\lfloor \frac{Mr_x}{N} \rfloor$ bits to all entries, and then allocate another extra bit to the first $Mr_x - \lfloor \frac{Mr_x}{N} \rfloor N$ entries of \mathbf{Y} , where $\lfloor \cdot \rfloor$ denotes the floor operator. A similar approach is employed for the direct and support-set coding schemes. Furthermore, using the direct coding (or, support set coding), if $r_x < 1$ bit (or, $r_x < \frac{K}{M} \log_2 M + \frac{K}{M}$ bit), we quantize the first Mr_x (or, $K \log_2 M + K$) coefficients of the reconstructed signal $\bar{\mathbf{x}}$ by 1 bit, and the remaining coefficients at the decoder are filled with zero.

In order to offer insights into the efficiency of our proposed AbS quantization algorithms, we compare their performances with that of the (optimal) joint quantization (8) which provides a benchmark performance. As mentioned earlier, the complexity of the joint quantization scheme (8) grows exponentially with the total bit budget. Due to this exponential complexity, a benchmark performance evaluation is impossible to realize at high rate and dimension. However, in order to gain insights regarding the performance of our AbS-based quantizers, we practically implemented the joint optimization method (8) at very low quantization bit-rate and dimension. The curves in Figure 3 illustrate the performance (NMSE) as a function of measurement rate α for the simulation setup $M = 40$, $K = 3$ and $r_x = 0.5$ bit per source component. As can be seen, while the nearest-neighbor coding gives a poor performance, the AbS-based quantizers perform nearly optimal, and the gap between the performance of the non-sequential/sequential AbS quantizer and the benchmark joint quantizer is less than 0.5 dB/1 dB in the worst case (at $\alpha = 0.45$).

Using a larger simulation parameter set $M = 512$, $K = 35$ (sparsity ratio $\approx 6.8\%$), we illustrate the performance (NMSE) of the quantization algorithms as a function of measurement

rate α , shown in Figure 4 and Figure 5 for fixed quantization bit-rates $r_x = 0.5$ and $r_x = 0.75$ bit per component of \mathbf{X} , respectively. First, let us interpret the performance behavior of the schemes for quantization in the CS measurement domain, i.e., nearest neighbor coding and the proposed AbS quantization schemes. It is worth pointing out that increasing the measurement rate α , given a fixed bit budget r_x , has two different effects on the performance. On one hand, it facilitates a more precise reconstruction both at the encoder and the decoder due to increasing number of measurements. On the other hand, it reduces (increases) quantization bit-rate (quantization noise) per entry of the measurement vector since $r_y = r_x/\alpha$. Following these facts, it can be observed from the curves that given a very small values of α , the sparse reconstruction algorithm fails to detect the sparsity pattern and reconstruct the source. This results in a poor performance although the quantization bit-rate per entry is high. As α increases to a certain amount, the reconstruction algorithm succeeds to reconstruct the sparse source precisely out of the measurements since the number of measurements is sufficient, and the quantization noise is small enough. At this point ($\alpha = 0.25$ for Figure 4 and Figure 5), the curves reach the best performance. However, for higher α 's, due to the limited quantization bit-rate r_y , the quantization noise per entry increases which leads to a poorer performance. Among these schemes, the performance of the nearest-neighbor coding is the worst since it does not take the end-to-end distortion into account. However, the sequential and non-sequential AbS-based algorithms improve the performance significantly by exploiting the AbS framework. The gap between the performance of the two AbS schemes reduces as α increases since CS measurements tend to become i.i.d. RV's, and therefore, the non-sequential optimization method does not provide any extra gain. Observing the curves in Figure 4 and Figure 5, it can be also found that the MSE-minimizing measurement rate using both $r_x = 0.5$ and $r_x = 0.75$ occurs at $\alpha = 0.25$. We cannot, in general, claim that the optimal α would be the same at all quantization bit-rate regions. However, it can be inferred that the curves reach their minima, and then they take an upward trend.

Next, we evaluate the performance of the schemes for quantization in the reconstructed signal domain. At small α 's, the reconstruction algorithm fails to reconstruct the locally sparse source. It can be seen that as α increases, and the sparse reconstruction algorithm is able to reconstruct the input signal vector, the curve does not vary much by further increasing measurement rate since the allocated quantization bits using this scheme are independent of number of measurements, unlike the quantization schemes for measurement domain. The direct coding leads to a poor performance since this scheme does not take into consideration the sparsity pattern, while the support set coding improves the performance by only quantizing the magnitude of the non-zero coefficients in the estimated support set. By adaptively choosing the better performance among the quantization schemes in signal and CS measurement domains, one can benefit from both schemes at all ranges of α which is labeled by adaptive coding. However, note that in the spirit of exploiting CS for practical

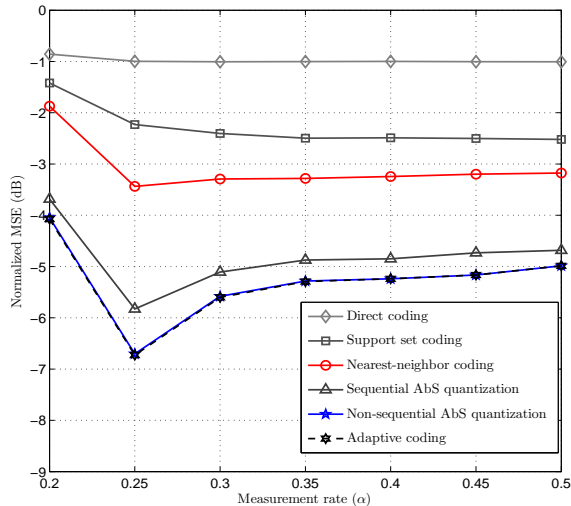


Fig. 4. Normalized MSE (in dB) as a function of measurement rate ($\alpha = N/M$) given a fixed quantization bit-rate using different quantization schemes. The parameters are chosen as $M = 512$, $K = 35$ and $r_x = 0.5$ bit per component of the input vector which is equivalent to total bit budget $R_x = 256$ bits.

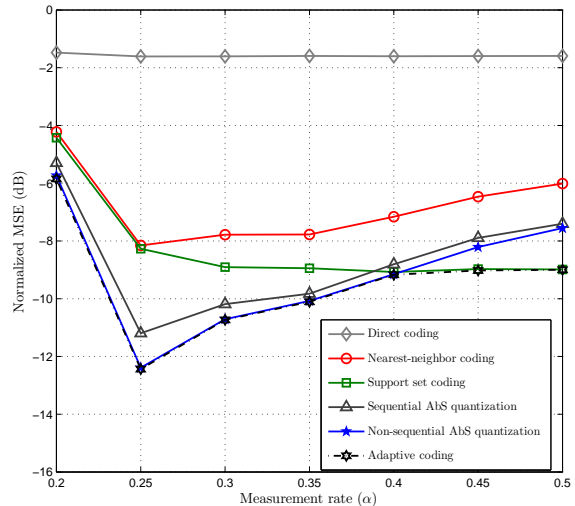


Fig. 5. Normalized MSE (in dB) as a function of measurement rate ($\alpha = N/M$) given a fixed quantization bit-rate using different quantization schemes. The parameters are chosen as $M = 512$, $K = 35$ and $r_x = 0.75$ bit per component of the input vector which is equivalent to total bit budget $R_x = 384$ bits.

applications, we are mainly interested in the lower ranges of α , for example, at $\alpha = 0.25$ in Figure 5, where the AbS quantization schemes achieve at least a considerable 3 dB reduction in MSE compared to other competing schemes.

In addition to the experimental results presented in Figure 4 and Figure 5, we also carried out similar simulations with the parameter set $M = 1000$, $K = 40$ (sparsity ratio = 4%), $r_x = 0.4$ bit per component of \mathbf{X} . We found similar trend in performances of all competing schemes and that the AbS quantization algorithms improve the performance at least 3 dB compared to the nearest-neighbor coding. Nevertheless, the direct and support set coding schemes achieve poor performances.

Next, we examine the performance (NMSE) as a function of quantization bit-rate per entry of \mathbf{X} , i.e. r_x , which is reported for the simulation setup $M = 512$, $K = 35$ at fixed $\alpha = 0.25$ in Figure 6. It can be observed that in low and moderate bit-rate regimes, the AbS quantization outperforms the other schemes, while at high bit-rates the support set coding attains a slightly better performance. Using the adaptive coding scheme, one can achieve the best performance among the competing schemes at all ranges of quantization bit-rate. At very high bit-rates, it is observed that the NMSE curves saturate. This is due to the fact that although the distortion due to quantization decreases as quantization bit-rate increases, the distortion due to sparse reconstruction (because of low number of measurements) still exists at a fixed measurement rate. At very high bit-rates, in order to eliminate the MSE floor, we need more number of measurements (sensors) so that the sparse reconstruction distortion becomes negligible. However, note that the use of more number measurements is restricted in CS.

C. Experimental Results: Comparison with Existing Schemes

In order to verify the efficiency of our proposed AbS schemes, we compare the performance of the sequential AbS quantizer vis-a-vis LASSO-optimized quantizer and uniform quantizer of [14]. It is note-worthy while our design method is neither asymptotic nor limited to any particular sparse reconstruction algorithm, the LASSO-optimized quantizer of [14] is based on two main assumptions: (1) asymptotic quantization bit-rate, (2) LASSO reconstruction for recovering a sparse source from quantized measurements. Another important design difference between the LASSO-optimized quantizer and our AbS-based quantizer schemes is described as follows. The LASSO-optimized scheme is based on the state-of-the-art distributed functional scalar quantizer (DFSQ) [34] for quantization of a scalar function. Therefore, it only minimizes the MSE of a scalar function (arbitrary output of the LASSO sparse reconstruction) of the measurement vector. However, in our AbS schemes, we design the quantizers by taking into consideration the MSE between a source vector and its final reconstruction vector.

In order to reconstruct a sparse source $\hat{\mathbf{x}}$ from quantized measurements $\hat{\mathbf{y}}$, the LASSO reconstruction algorithm aims at solving the following convex optimization problem

$$\hat{\mathbf{x}}_{\text{LASSO}} = \underset{\mathbf{x} \in \mathbb{R}^M}{\text{argmin}} \|\hat{\mathbf{y}} - \Phi \mathbf{x}\|_2 + \mu \|\mathbf{x}\|_1, \quad (15)$$

where $\mu > 0$ is a fixed user parameter. We solve (15) using the SPGL1 toolbox [35].

For fair comparison, we use the same kind of sources as used in the schemes of [14]. We randomly generate the non-zero coefficients of \mathbf{X} from uniform distribution $\mathcal{U}[-1, 1]$, and use the LASSO reconstruction (15) with $\mu = 10^{-3}$ (the choice of μ is experimentally verified to achieve the best

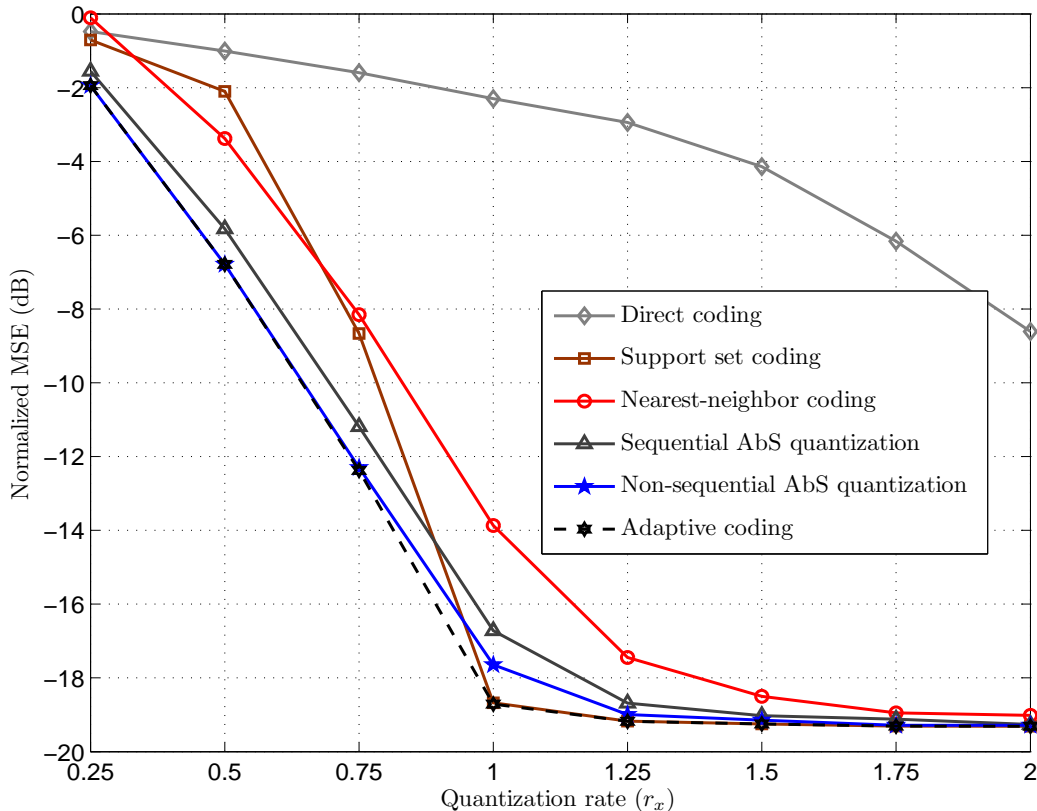


Fig. 6. Normalized MSE (in dB) as a function of rate per each component of \mathbf{X} (i.e., r_x) given a fixed measurement rate using different quantization schemes. The parameters are chosen as $M = 512$, $K = 35$ and $\alpha = 0.25$. The values on the x-axis are equivalent to $r_y = 1$ bit to $r_y = 8$ bits per measurement entry. At very high rates, the curves achieve MSE floor due to the fixed sparse reconstruction distortion.

performance). All other simulation setups are the same as those given in Section V-A. Also, simulation parameters are set to $M = 100$, $K = 10$ and $\alpha = 0.5$. In Figure 7, the performance (in terms of NMSE) as a function of quantization bit-rate r_x is illustrated for the sequential AbS quantizer as well as LASSO-optimized quantizer and uniform quantizer design schemes. As would be expected, the uniform quantizer provides a poor performance because of inappropriate codepoints for quantization of measurement entries. It can be also seen that the proposed sequential AbS quantization gives a better performance (almost 3 dB) than the LASSO-optimized quantizer.

VI. CONCLUSIONS

Due to non-linear behavior of sparse reconstructions, the effect of quantization on CS measurements would reflect in a non-linear manner in the signal reconstruction. To handle the non-linearity, we have developed AbS-based quantization schemes, and shown that a significant improvement in performance can be achieved. We have found that the AbS schemes have a limitation at high quantization bit-rates, and hence we have developed an adaptive coding scheme suited for all scenarios. Furthermore, comparisons with existing quantiza-

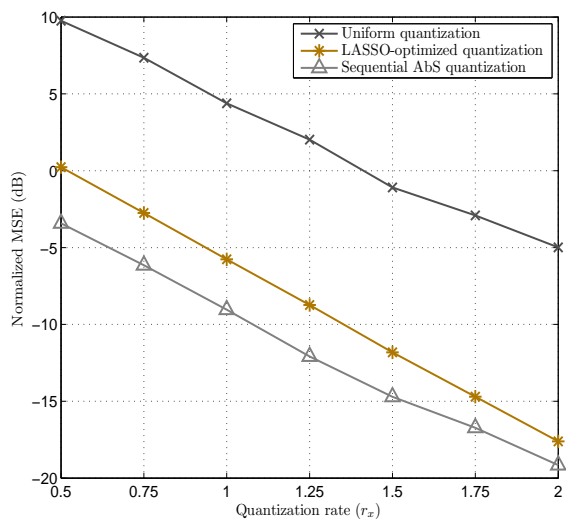


Fig. 7. Normalized MSE (in dB) as a function of rate per each component of \mathbf{X} (i.e., r_x) given a fixed measurement rate $\alpha = 0.5$ using the proposed sequential AbS, LASSO-optimized and uniform quantizer schemes of [14]. The parameters are chosen as $M = 100$, $K = 10$ and $\alpha = 0.5$.

tion algorithms for CS measurements have demonstrated the efficiency of our AbS-based quantization schemes.

APPENDIX

We briefly describe the orthogonal matching pursuit (OMP) algorithm [22], [23] for reconstructing a sparse source from quantized CS measurement vector $\hat{\mathbf{y}}$, where the sensing matrix and sparsity level are provided as well. The low-complexity OMP scheme is an iterative algorithm where it performs a matched filter operation and an orthogonal projection at each iteration. Using pseudo-inverse matrix inversion, the orthogonal projection operations can be performed recursively. The main steps of the OMP are summarized in Algorithm 2.

Algorithm 2 : Orthogonal matching Pursuit (OMP)

```

1: input:  $\Phi, \hat{\mathbf{y}} = \hat{\mathbf{y}}, K$ 
2: initialization:  $l \leftarrow 0$  (Iteration counter)  $\mathbf{r}^{(0)} \leftarrow \hat{\mathbf{y}}$  (Initial residual),  $\mathcal{S}^{(0)} \leftarrow \emptyset$  (Initial support set)
3: repeat
4:    $l \leftarrow l + 1$ 
5:    $i^{(l)} \leftarrow \text{index of the highest amplitude of } \Phi^T \mathbf{r}^{(l-1)}$ 
6:    $\mathcal{S}^{(l)} \leftarrow \mathcal{S}^{(l-1)} \cup i^{(l)}$  (Note:  $|\mathcal{S}^{(l)}| = l$ )
7:    $\mathbf{r}^{(l)} \leftarrow \hat{\mathbf{y}} - \Phi_{\mathcal{S}^{(l)}} \Phi_{\mathcal{S}^{(l)}}^\dagger \hat{\mathbf{y}}$  (Orthogonal projection)
8: until ( $(\|\mathbf{r}^{(l)}\|_2 > \|\mathbf{r}^{(l-1)}\|_2)$  or  $(l > K)$ )
9:  $l \leftarrow l - 1$  (Previous iteration)
10: output:  $\hat{\mathbf{x}} \in \mathbb{R}^M$ , satisfying  $\hat{\mathbf{x}}_{\mathcal{S}^{(l)}} = \Phi_{\mathcal{S}^{(l)}}^\dagger \hat{\mathbf{y}}$  and  $\hat{\mathbf{x}}_{\bar{\mathcal{S}}^{(l)}} = \mathbf{0}$ 

```

REFERENCES

- [1] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Sig. Proc. Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [2] A. Yang, M. Gastpar, R. Bajcsy, and S. Sastry, "Distributed sensor perception via sparse representation," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1077–1088, June 2010.
- [3] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [4] C. Güntürk, M. Lammers, A. Powell, R. Saab, and O. Yılmaz, "Sigma delta quantization for compressed sensing," in *Annual Conference Inf. Sciences and Systems*, March 2010, pp. 1–6.
- [5] A. Zymnis, S. Boyd, and E. Candes, "Compressed sensing with quantized measurements," *IEEE Sig. Proc. Lett.*, vol. 17, no. 2, pp. 149–152, Feb. 2010.
- [6] W. Dai and O. Milenkovic, "Information theoretical and algorithmic approaches to quantized compressive sensing," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1857–1866, July 2011.
- [7] L. Jacques, D. Hammond, and J. Fadili, "Dequantizing compressed sensing: When oversampling and non-Gaussian constraints combine," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 559–571, Jan. 2011.
- [8] M. Yan, Y. Yang, and S. Osher, "Robust 1-bit compressive sensing using adaptive outlier pursuit," *IEEE Trans. Sig. Proc.*, vol. 60, no. 7, pp. 3868–3875, July 2012.
- [9] L. Jacques, J. Laska, P. Boufounos, and R. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2082–2102, 2013.
- [10] U. S. Kamilov, V. K. Goyal, and S. Rangan, "Message-passing dequantization with applications to compressed sensing," *IEEE Trans. Sig. Proc.*, vol. 60, no. 12, pp. 6270–6281, Dec. 2012.
- [11] V. Goyal, A. Fletcher, and S. Rangan, "Compressive sampling and lossy compression," *IEEE Sig. Proc. Mag.*, vol. 25, no. 2, pp. 48–56, March 2008.
- [12] J. Laska and R. Baraniuk, "Regime change: Bit-depth versus measurement-rate in compressive sensing," *IEEE Trans. Sig. Proc.*, vol. 60, no. 7, pp. 3496–3505, July 2012.
- [13] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Royal Stat. Soc.*, vol. 58, no. 1, pp. 267–288, Ser. B 1996.
- [14] J. Sun and V. Goyal, "Optimal quantization of random measurements in compressed sensing," in *IEEE Int. Symp. Inf. Theory*, July 2009, pp. 6–10.
- [15] U. Kamilov, V. Goyal, and S. Rangan, "Optimal quantization for compressive sensing under message passing reconstruction," in *IEEE Int. Symp. Inf. Theory*, 31 2011–Aug. 5 2011, pp. 459–463.
- [16] P. Boufounos, "Universal rate-efficient scalar quantization," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1861–1872, March 2012.
- [17] P. Kroon and E. Deprettere, "A class of analysis-by-synthesis predictive coders for high quality speech coding at rates between 4.8 and 16 kbit/s," *IEEE Journal Select. Areas Commun.*, vol. 6, no. 2, pp. 353–363, Feb 1988.
- [18] K. Aizawa and T. Huang, "Model-based image coding advanced video coding techniques for very low bit-rate applications," *Proceedings of the IEEE*, vol. 83, no. 2, pp. 259–271, Feb. 1995.
- [19] E. George and M. Smith, "Speech analysis/synthesis and modification using an analysis-by-synthesis/overlap-add sinusoidal model," *IEEE Trans. Speech Audio Proc.*, vol. 5, no. 5, pp. 389–406, Sep. 1997.
- [20] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [21] E. Candes and T. Tao, "Rejoinder: the Dantzig selector: statistical estimation when p is much larger than n ," *Annals of Statistics*, vol. 35, pp. 2392–2404, 2007.
- [22] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [23] T. Blumensath and M. Davies, "Gradient pursuits," *IEEE Trans. Sig. Proc.*, vol. 56, no. 6, pp. 2370–2382, June 2008.
- [24] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
- [25] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.
- [26] S. Chatterjee, D. Sundman, M. Vehkaperä, and M. Skoglund, "Projection-based and look-ahead strategies for atom selection," *IEEE Trans. Sig. Proc.*, vol. 60, no. 2, pp. 634–647, Feb. 2012.
- [27] E. G. Larsson and Y. Selen, "Linear regression with a sparse parameter vector," *IEEE Trans. Sig. Proc.*, vol. 55, no. 2, pp. 451–460, Feb. 2007.
- [28] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Sig. Proc.*, vol. 56, no. 6, pp. 2346–2356, June 2008.
- [29] M. Elad and I. Yavneh, "A plurality of sparse representations is better than the sparsest one alone," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4701–4714, Oct. 2009.
- [30] M. Protter, I. Yavneh, and M. Elad, "Closed-form MMSE estimation for signal denoising under sparse representation modeling over a unitary dictionary," *IEEE Trans. Sig. Proc.*, vol. 58, no. 7, pp. 3471–3484, July 2010.
- [31] K. Qiu and A. Dogandic, "Sparse signal reconstruction via ECME hard thresholding," *IEEE Trans. Sig. Proc.*, vol. 60, no. 9, pp. 4551–4569, Sep. 2012.
- [32] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1991.
- [33] J. N. Laska, P. T. Boufounos, M. A. Davenport, and R. G. Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing," *Applied and Computational Harmonic Analysis*, vol. 31, no. 3, pp. 429–443, 2011.
- [34] V. Misra, V. Goyal, and L. Varshney, "Distributed scalar quantization for computing: High-resolution analysis and extensions," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5298–5325, 2011.
- [35] E. van den Berg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," June 2007, <http://www.cs.ubc.ca/labs/scl/spgl1>.